PACKET LOSS COMPENSATION METHOD USING INJECTION OF SPECTRALLY SHAPED NOISE

FIELD OF THE INVENTION

This invention relates in general to packetized voice communication systems, and more particularly to a method of compensating for lost packets in a packetized voice system by injecting spectrally shaped noise.

BACKGROUND OF THE INVENTION

Transmission of voice over packet networks has emerged in recent years as a replacement for traditional legacy PBX systems for telephone communications. A packetized voice transmission system comprises a transmitter and a receiver. The transmitter collects voice samples and groups them into packets for transmission across a network to the receiver. The data itself may be companded according to u-law or A-law, as defined in ITU-T specification G.711. Other companding/vocoding techniques, such as G.729, G.723.1, can also be used.

When using a packet based network, packet losses due to congestion in the network can produce significant degradation of the performance of echo cancellers. The effects introduced by packet loss depend to a large extent on the techniques used to recover lost packets. Packet loss recovery techniques can be divided into two classes: sender-based repair and receiver-based repair [see C. Perkins, O. Hodson and V. Hardman, "A Survey of Packet Loss Recovery Techniques for Streaming Audio," IEEE Network, Sept./Oct. 1998, pp. 40-48]. Receiver-based repair is also referred to in the art as error concealment.

Among known error concealment techniques, those based on packet insertion have found popularity due to ease of implementation. According to such insertion-based recovery techniques a replacement packet is inserted to fill the gap left by a lost packet. The replacement packet can be one of either silence, white noise or repetition of the previous packet. Silence substitution is simple to implement but performs poorly. Since silence substitution fills the gap left by a lost packet with silence in order to maintain the timing

relationship between the surrounding packets, the performance of silence substitution degrades rapidly as packet sizes increases, and quality is unacceptably bad for the 40 ms packet size in common use in network audio conferencing tools. Some studies have shown that inserting white noise, instead of silence, can improve intelligibility [see G. A. Miller and J. C. R. Licklider, "The Intelligibility of Interrupted Speech," *J. Acoust. Soc. Amer.*, vol. 22, no. 2, 1950, pp. 167-73; and R. M. Warren, *Auditory Perception*, Pergamon Press, 1982]. Among the three methods of packet insertion, repetition of the previous packet gives best voice quality due to the similarity between the neighboring voice segments.

Although the uses of white noise and previous packets may yield better speech quality than silence substitution does, these techniques interfere with proper operation of network echo cancellers. The substitution of white noise results in a sudden change in the spectral characteristics of the signal, causing severe degradation of echo return loss enhancement (ERLE). When substituting a previous packet, the fill-in packet is the same as the previous packet, which means that the two packets are highly correlated. This reduces the convergence rate and results in slow recovery from the packet loss.

SUMMARY OF THE INVENTION

According to the present invention, a new insertion-based error concealment method and apparatus are provided whereby, instead of directly inserting white noise, a filter is created to shape the white noise. The filtered white noise is then used to replace lost data. The method of the present invention is implemented by first estimating the power spectrum of the previous frame; then designing a filter with transfer function H(f), where $|H(f)|^2$=the estimated power spectrum; and finally generating the replacement packet using noise which has been spectrally modified by the filter. The resulting filtered noise has the same power spectrum as the previous packet but is not highly correlated with it.

BRIEF DESCRIPTION OF THE DRAWINGS

A detailed description of a preferred embodiment of the present invention is provided herein below with reference to the drawings in which:

Figure 1 is a block diagram showing a lost packet generator for use in a data packet transmission system according to the present invention;

Figure 2 is a flowchart showing steps in the lost packet compensation method of the present invention; and

Figure 3 is a graph showing a comparison of the impact of packet loss compensation on ERLE using the method and apparatus of the present invention with the prior art.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

With reference to Figures 1 and 2, a new apparatus and method are shown according to the preferred embodiment, for packet loss compensation in a voice communication system. A buffer 3 receives and stores successive frames of received voice data. A packet loss detector 5 detects lost packets and in response operates a pair of switches 7 and 9, as discussed in greater detail below. The design and operation of buffer 3 and packet loss detector 5 will be well known to a person of ordinary skill in the art and are not, therefore, discussed in further detail herein.

In response to detecting a lost packet, switch 7 closes and the previous voice packet stored in buffer 3 is applied to power spectrum estimator 11. Power estimator 11 implements Welch's averaged periodogram method for estimating the power signal $\mathbf{P}(\omega)$, (see P. D. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra", IEEE Trans. Audio Elecrtoacoust., Vol AU-15, June 1970, pp. 70 – 73), although any spectral estimation algorithm will suffice. The output of the spectrum estimator is sent to a filter coefficients calculator 13. The filter coefficients calculator 13 designs an FFT filter 15 with transfer function H(f), where $|H(f)|^2$=the estimated power spectrum. filter coefficients calculator 13 and filter 15 may be implemented using a digital signal processor (DSP) using well known techniques. According to a successful implementation a 64 bit FFT was used. White noise is output from generator 17 to the filter 15 so that the shapes the white noise to the characteristics of the voice signal. As indicated above, packet loss detector 5 operates switch 9 so that in response to a lost packet, the filtered noise from filter 15 is output to replace lost data. The filtered noise has the same power spectrum as the previous frame. Due

to the similarity between the neighboring frames, the filtered noise is more similar to the lost packet than unfiltered white noise is.

Figure 3 shows the comparative ERLE performance of the lost packet compensation method of the present invention relative to other techniques. It can be seen that inserting silence and white noise exhibit the smallest and greatest impact on the ERLE performance, respectively. However, the degradation of ERLE is smaller using the system according to the present invention than when using substitution of white noise, and the impact on ERLE decays quicker compared to the substitution of previous packets.

Alternative embodiments and variations of the invention are possible. For example, although the inventive method and apparatus have been described in terms of voice transmission over IP networks, it is contemplated that the principles of the invention may be extended to other asynchronous systems such as ATM networks. Also, whereas the preferred embodiment sets forth the use of Welch's algorithm and an FFT filter for spectral estimation and filtering, respectively, it is possible to use other spectral estimation algorithms (e.g. Linear Predictive Coding (LPC)), and other filtering (e.g. using LPC coefficients).

All such changes and modifications may be made without departing from the sphere and scope of the invention as defined by the claims appended hereto.